┌───┐
│ This Raw Listing contains the General   │
│ Information Section and up to the first 5 pages. │
└───┘

```
 1                                   SEQUENCE LISTING       ENTERED
 2
 3    (1)    General Information:
 4
 5      (i) APPLICANT: Peter DORMER
 6
 7     (ii) TITLE OF INVENTION: PROTEIN WITH DIFFERENTIATION-INDUCING
 8                    ACTIVITY FOR FRIEND'S ERYTHROLEUKEMIA CELL LINES
 9
10    (iii) NUMBER OF SEQUENCES: 10
11
12     (iv) CORRESPONDENCE ADDRESS:
13            (A) ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER
14            (B) STREET: 99 Canal Center Plaza, Suite 300
15            (C) CITY: Alexandria
16            (D) STATE: VA
17            (E) COUNTRY: USA
18            (F) ZIP: 22314
19
20      (v) COMPUTER READABLE FORM:
21            (A) MEDIUM TYPE: Floppy disk
22            (B) COMPUTER: IBM PC compatible
23            (C) OPERATING SYSTEM: PC-DOS/MS-DOS
24            (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
25
26     (vi) CURRENT APPLICATION DATA:
27            (A) APPLICATION NUMBER:
28            (B) FILING DATE:
29            (C) CLASSIFICATION:
30
31   (viii) ATTORNEY/AGENT INFORMATION:
32            (A) NAME: Presta, Frank P.
33            (B) REGISTRATION NUMBER: 19,828
34            (C) REFERENCE/DOCKET NUMBER: 3428-005
35
36     (ix) TELECOMMUNICATION INFORMATION:
37            (A) TELEPHONE: (703) 684-1111
38            (B) TELEFAX: (703) 684-1124
39
40
41
42    (2) INFORMATION FOR SEQ ID NO: 1:
43
44      (i) SEQUENCE CHARACTERISTICS:
45            (A) LENGTH: 1495 base pairs
46            (B) TYPE: nucleic acid
```

```
47                  (C) STRANDEDNESS: single
48                  (D) TOPOLOGY: linear
49
50       (ii) MOLECULE TYPE: cDNA to mRNA
51
52      (iii) HYPOTHETICAL: YES
53
54       (iv) ANTI-SENSE: NO
55
56       (vi) ORIGINAL SOURCE:
57                  (A) ORGANISM: Mus musculus
58
59
60
61       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
62
63   CCGACCGTGC GGACTTAAGA TGGAGGCACT TCCTGTCTGC GGCGGGAAGA GAAGGCTCGG        60
64
65   TCGGAGCCGG GAATGCTGGG ACTTGTAGTG CGTAGTCAAT GGTTCTCTAT GGGCTTTCAG       120
66
67   AGTGAGTGGC GGGAAGGCGG CCCCGAGGCA TGCTGGGAGT TGTAGTCCTG CCGTCGTCAA       180
68
69   TGGTTCTCTA TGGGCTTTCA GAGTGAGTGG CGGGAAGGCG GCCCCGAGGC ATGCTGGGAG       240
70
71   TTGTAGTCCT GCCATAGTCA ATGGTTCTCT ATGGGCTTTC AGACTGAGTG GCGGGAAGGC       300
72
73   GGCCCCGAGG CATGCTGGGA GTTGCAGCGC CATGTTTTAA AGCACGCGTT TCTCTGTATA       360
74
75   GACCTGGCTG TGGATTTTTC GCTAATTCTT TTTTTAGCT TTATTTTTAA TTTTTACTTT        420
76
77   TTCACACAGG ATTTCTCTTT ATAGCCTTGG CTACCGTTTT TTCCCTAATT ATTCTCCTTT       480
78
79   TCATTTTGGT TTATTTTTTT TTAATTTTGG TTTTTTTAAG ACAGGGTTTC TCTGTATAGA       540
80
81   CCTGGCTGTG GATTTCTCAC TAATTATTTT TTTTAGCTTT ATTTTTAATT TTTACTTTTT       600
82
83   CACACAGGAT TTCTCTTTAT AGCCTTGGCT ACCGTTTTTT CCGTAATTAT TCTTATTTTC       660
84
85   ATTTTGGTTT ATTTTTTAAT TTTAATTTTT GATTTTGGAG ACAGGGTTTC TCTTTTAGCC       720
86
87   GCAGCTATGG TTTCTGCCCT AATTATTCTT GTCCTTATTT GTAATTTAAT TCTTAATTTA       780
88
89   ATTTAATTTA TAATTTTGTT GTAAGTTTTT CTGTGGGCGT GAATGGAAAG TCTAACCCGT       840
90
91   GTTTCTCTGT TCAGCGTCCG CCGGTCACGG CCGCCGCCCC CAGCGACGTC ACCCACACGC       900
92
93   GCAGAAGCGG ACGCCGCGGT CAAGATGTCT CTGCCATGCC CACGGGACGC ACGGACGCAC       960
94
95   GGACGGACGG ACGGACTCCA CAAGGTAGGA AGCCTGCGCC GACCGCACCG CCGCACCCAC      1020
96
97   CACAGCACAC AGGACACACG CGGGCCCCGC GCCCGCCCAG GCACACGCGG CACACACGGC      1080
98
99   ACACACGGCA GGCAGGCCAG GCACACGCAT CCGCAGGACC CGCCGCACCC GCCACGCAGA      1140
```

```
100
101   CACGGACGAG CCGCCGCGGT CAAGATGTTC ACCCGCCGCG GTCAAGATGT ATGTGCCACC      1200
102
103   GACCCTCGCC CCGCTGGACG GACGGACGGA CGCACGCACG CCGTCAGCGT CCACCGGTCA      1260
104
105   CTGCCGCCGC CCACAGTGAT GTCACCCACG AAAGCACACA CGTAGAAGCG GACGCCGTGG      1320
106
107   TCAAGATGTC TCTGCCATCC CCACAGGACG GACGGACGGA CTCCACAAGG TGCGCGTGTC      1380
108
109   GCCGAGGCCG CCAGGACGGA GCGATTCTCA CGGAGGAAGG AGCACGCCAA CAGGGCCTGA      1440
110
111   CTGCGTACAG ACATGTCCCC CTCAATAAAA TTGCAGTTGA AATGGAAAAA AAAAA           1495
112
113   (2) INFORMATION FOR SEQ ID NO: 2:
114
115        (i) SEQUENCE CHARACTERISTICS:
116             (A) LENGTH: 715 base pairs
117             (B) TYPE: nucleic acid
118             (C) STRANDEDNESS: single
119             (D) TOPOLOGY: linear
120
121     (ii) MOLECULE TYPE: cDNA to mRNA
122
123    (iii) HYPOTHETICAL: YES
124
125
126     (ix) FEATURE:
127             (A) NAME/KEY: CDS
128             (B) LOCATION:155..688
129
130
131     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
132
133   CGCGCCCGCC CGGGATCCCC AGCTGCCGCC GCGCCCGCCC GCCCGCCCGG GGCCCCCGCT        60
134
135   GCAGAACCGT GACCGTCCGC CGGTCACGGC CGCCGCCCCC AGCGACGTCA CCCACACGCG       120
136
137   CAGAAGCGGA CGCCGCGGTC AAGATGTCTC TGCC ATG CCC ACG GGA CGC ACG          172
138                                          Met Pro Thr Gly Arg Thr
139                                           1                   5
140
141   GAC GCA CGG ACG GAC GGA CTG ACT CCA CAA GGT AGG AAG CCT GCG CCG        220
142   Asp Ala Arg Thr Asp Gly Leu Thr Pro Gln Gly Arg Lys Pro Ala Pro
143               10              15                  20
144
145   ACC GCA CCG CCG CAC CCA CCA CAG CAC ACA GGA CAC ACG CGG GCC CCG        268
146   Thr Ala Pro Pro His Pro Pro Gln His Thr Gly His Thr Arg Ala Pro
147           25              30                  35
148
149   CGC CCG CCC AGG CAC ACG CGG CAC ACA CGG CAC ACA CGG CAG GCA GGC        316
150   Arg Pro Pro Arg His Thr Arg His Thr Arg His Thr Arg Gln Ala Gly
151       40              45                  50
152
```

INPUT SET: S25841.raw

```
153   CAG GCA CAC GCA TCC GCA GGA CCC GCC GCA CCC GCC ACG CAG ACA CGG        364
154   Gln Ala His Ala Ser Ala Gly Pro Ala Ala Pro Ala Thr Gln Thr Arg
155    55              60              65                  70
156
157   ACG AGC CGC CGC GGT CAA GAT GTT CAC CCG CCG CGG TCA AGA TGT ATG        412
158   Thr Ser Arg Arg Gly Gln Asp Val His Pro Pro Arg Ser Arg Cys Met
159                    75              80              85
160
161   TGC CAC CGA CCC TCG CCC CGC TGG ACG GAC GGA CGG ACG CGC GCA CGC        460
162   Cys His Arg Pro Ser Pro Arg Trp Thr Asp Gly Arg Thr Arg Ala Arg
163                90              95              100
164
165   CGT CAG CGT CCA CCG GTC ACT GCC GCC GCC CAC AGT GAC GTC ACC CAC        508
166   Arg Gln Arg Pro Pro Val Thr Ala Ala Ala His Ser Asp Val Thr His
167           105             110             115
168
169   GAA AGC ACA CAC GTA GAA GCG GAC GCC GTG GTC AAG ATG TCT CTG CCA        556
170   Glu Ser Thr His Val Glu Ala Asp Ala Val Val Lys Met Ser Leu Pro
171       120             125             130
172
173   TCC CCA CAG GAC GGA CGG ACG GAC TCC ACA AGG TGC GCG TGT CGC CGA        604
174   Ser Pro Gln Asp Gly Arg Thr Asp Ser Thr Arg Cys Ala Cys Arg Arg
175   135             140             145                 150
176
177   GGC CGC CAG GAT GGA GCG ATT CTC ACG GAG GAA GGA GCA CGC CAA CAG        652
178   Gly Arg Gln Asp Gly Ala Ile Leu Thr Glu Glu Gly Ala Arg Gln Gln
179                155             160             165
180
181   GGC CTG ACT GCG TAC AGA AAT GCC CCC CCT CAA TAA AATTGCAGTT            698
182   Gly Leu Thr Ala Tyr Arg Asn Ala Pro Pro Gln  *
183                170             175
184
185   GAAATGGAAA AAAAAAA                                                     715
186
187
188   (2) INFORMATION FOR SEQ ID NO: 3:
189
190        (i) SEQUENCE CHARACTERISTICS:
191            (A) LENGTH: 177 amino acids
192            (B) TYPE: amino acid
193            (D) TOPOLOGY: linear
194
195     (ii) MOLECULE TYPE: protein
196     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
197
198   Met Pro Thr Gly Arg Thr Asp Ala Arg Thr Asp Gly Leu Thr Pro Gln
199    1               5                   10                  15
200
201   Gly Arg Lys Pro Ala Pro Thr Ala Pro Pro His Pro Pro Gln His Thr
202                20              25                  30
203
204   Gly His Thr Arg Ala Pro Arg Pro Pro Arg His Thr Arg His Thr Arg
205                35              40                  45
```

INPUT SET: S25841.raw

```
206
207    His Thr Arg Gln Ala Gly Gln Ala His Ala Ser Ala Gly Pro Ala Ala
208         50              55              60
209
210    Pro Ala Thr Gln Thr Arg Thr Ser Arg Arg Gly Gln Asp Val His Pro
211     65              70              75              80
212
213    Pro Arg Ser Arg Cys Met Cys His Arg Pro Ser Pro Arg Trp Thr Asp
214                 85              90              95
215
216    Gly Arg Thr Arg Ala Arg Arg Gln Arg Pro Pro Val Thr Ala Ala Ala
217                 100             105             110
218
219    His Ser Asp Val Thr His Glu Ser Thr His Val Glu Ala Asp Ala Val
220         115             120             125
221
222    Val Lys Met Ser Leu Pro Ser Pro Gln Asp Gly Arg Thr Asp Ser Thr
223         130             135             140
224
225    Arg Cys Ala Cys Arg Arg Gly Arg Gln Asp Gly Ala Ile Leu Thr Glu
226    145             150             155             160
227
228    Glu Gly Ala Arg Gln Gln Gly Leu Thr Ala Tyr Arg Asn Ala Pro Pro
229                 165             170             175
230
231    Gln
232
233
234    (2) INFORMATION FOR SEQ ID NO: 4:
235
236        (i) SEQUENCE CHARACTERISTICS:
237            (A) LENGTH: 636 base pairs
238            (B) TYPE: nucleic acid
239            (C) STRANDEDNESS: single
240            (D) TOPOLOGY: linear
241
242     (ii) MOLECULE TYPE: cDNA to mRNA
243
244    (iii) HYPOTHETICAL: YES
245
246
247     (ix) FEATURE:
248            (A) NAME/KEY: CDS
249            (B) LOCATION:1..636
250
251
252     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
253
254    ATG GGG CTG CAG AAC CGT GAC CGT CCG CCG GTC ACG GCC GCC GCC CCC      48
255    Met Gly Leu Gln Asn Arg Asp Arg Pro Pro Val Thr Ala Ala Ala Pro
256     180             185             190
257
258    AGC GAC GTC ACC CAC ACG CGC AGA AGC GGA CGC CGC GGT CAA GAT GTC      96
```

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION   *US/08/973,303*

INPUT SET: S25841.raw

Line	Error	Original Text